hEART 2016 - 5th Symposium of the European Association for Research in Transportation, Delft University of Technology

Data-driven characterization of pedestrian flows

Marija Nikolić, Michel Bierlaire

September 14, 2016

Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

Motivation

CANDY
JEWLERS
FLOWERS
DAILY NEWS
VALENTINES DAY
GRILLMEISTER
ASSORTED
TRUFFLES

Background

Importance

- Understanding, reproducing and forecasting phenomena that characterize pedestrian traffic is necessary in order to provide services related to pedestrian safety and convenience

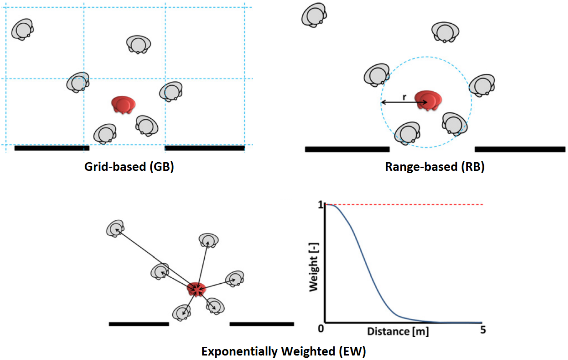
Indicators

- Density k (ped/m^2), speed v (m/s) and flow q (ped/ms)
- Used to observe and to model the flows of pedestrians
- Little concern is dedicated to the nature of spatial and temporal discretization underlying the definitions

Outline

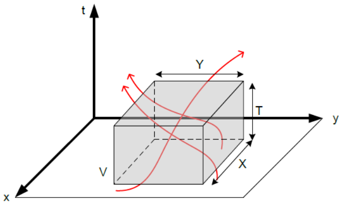
1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

Methods

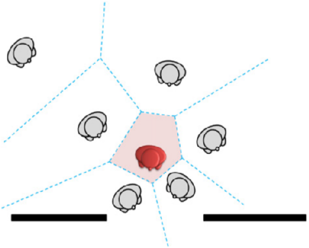


[Duives et al., 2015], [Helbing et al., 2007],[Steffen and Seyfried, 2010], [Saberi and Mahmassani, 2014],[van Wageningen-Kessels et al., 2014]

Methods



Edie (XY-T)



Voronoi-based (VB)

[Duives et al., 2015], [Helbing et al., 2007],[Steffen and Seyfried, 2010], [Saberi and Mahmassani, 2014],[van Wageningen-Kessels et al., 2014]

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Characteristics of methods

Method	Scale	Spatial aggregation		Temporal aggregation		Data type
		Unit	Assumptions	Unit	Assumptions	
XY-T	Macroscopic	Area	Shape Size Location	Interval	Duration	Trajectories
Grid-based (GB)	Macroscopic	Cell	Size Location	Interval	Duration	Trajectories Sync. sample
Range-based (RB)	Macroscopic	Circle	Radius Location	Interval	Duration	Trajectories Sync. sample
Exponentially-weighted (EW)	Macroscopic	Range	Influence function Range of influence	Interval	Duration	Trajectories Sync. sample
Voronoi-based (VB)	Microscopic	Voronoi cell	Boundary conditions	Interval	Duration	Trajectories Sync. sample

How to define the discretization...

...independent of arbitrary chosen values?

Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Data-driven approach

Keep calm and let data speak!

Preliminaries

- A space-time representation: $\Omega \subset \mathbb{R}^3$
- The distance along each of the two spatial axes is expressed in meters, and the unit for time is seconds
- $p = (p_x, p_y, p_t) = (x, y, t) \in \Omega$ represents a physical position (x, y) in space at a specific time t
- Assumption: Ω is convex (obstacle-free and bounded)
- Generator set Γ: pedestrian trajectories

$$\Gamma_i : \{p_i(t) | p_i(t) = (x_i(t), y_i(t), t)\}$$

$$\Gamma_i : \{p_{is} | p_{is} = (x_{is}, y_{is}, t_s)\}, t_s = [t_0, t_1, ..., t_f]$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

Data-driven discretization framework

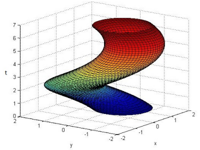
- 3D Voronoi diagrams associated with pedestrian trajectories
- Partitioning: the assignment of each point $p \in \Omega$ to one generator from Γ

$$\delta_\Gamma(p, \Gamma_i) = \begin{cases} 1, & D(p, \Gamma_i) \leq D(p, \Gamma_j), \forall j \neq i \\ 0, & \text{otherwise} \end{cases}$$

$$D(p, \Gamma_i) = \min_{p_i} \{ d(p, p_i) | p_i \in \Gamma_i, \Gamma_i \in \Gamma, p \in \Omega \}$$

- Discretization units: the set of points p assigned to the same generator

$$V_i = \{ p | \delta_\Gamma(p, \Gamma_i) = 1, p \in \Omega, \Gamma_i \in \Gamma \}$$



Data-driven discretization framework (cont.)

- The plane through the point $p_0 = (x_0, y_0, t_0)$ and with non-zero normal vector $\vec{n} = (a, b, c)$

$$\mathcal{P}_{\vec{n},p_0} : ax + by + ct + d = 0,$$

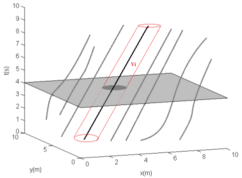
where $d = -ax_0 - by_0 - ct_0$

- The intersection of V_i and the plane $\mathcal{P}_{\vec{n},p_0}$

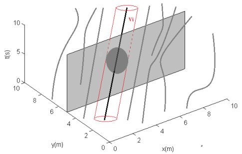
$$A(V_i, \mathcal{P}_{\vec{n},p_0}) = \{p | p \in \{V_i \cap \mathcal{P}_{\vec{n},p_0}\}\}$$

Data-driven discretization framework (cont.)

$A(V_i, \mathcal{P}_{(0,0,1),p_0}) =$
$\{p | p \in V_i \text{ and } p_t = t_0\}$



$A(V_i, \mathcal{P}_{(a,b,0),p_0}) =$
$\{p | p \in V_i \text{ and } ap_x + bp_y = ax_0 + by_0\}$



Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

Voronoi-based traffic indicators

Density indicator

$$k(x, y, t) = \frac{1}{|A(V_i, \mathcal{P}_{(0,0,1),(x,y,t)})|}$$

Flow indicator

$$\vec{q}_{(a,b,0)}(x, y, t) = \frac{1}{|A(V_i, \mathcal{P}_{(a,b,0),(x,y,t)})|}$$

Velocity indicator

$$\vec{v}_{(a,b,0)}(x, y, t) = \frac{\vec{q}_e(x,y,t)}{k(x,y,t)} = \frac{|A(V_i, \mathcal{P}_{(0,0,1),(x,y,t)})|}{|A(V_i, \mathcal{P}_{(a,b,0),(x,y,t)})|}$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Spatio-temporal distances

Spatial Euclidean distance (E-3DVoro)

$$d_E(p, p_i) = \begin{cases} \sqrt{(p - p_i)^T (p - p_i)}, & \Delta t = 0 \\ \infty, & \textit{otherwise} \end{cases}$$

Time-Transform distances (TT$_{\{1,2,3\}}$-3DVoro)

$$\mathrm{d}_{TT_1}(p, p_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + \alpha^2 (t - t_i)^2}$$

$$d_{TT_2}(p, p_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2} + \alpha_i(t_i)|(t - t_i)|$$

$$d_{TT_3}(p, p_i) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + \alpha_i^2(t_i)(t - t_i)^2}$$

α and α_i - conversion constants expressed in meters per second

Spatio-temporal distances

Predictive distance (P-3DVoro)

$$d_P(p, p_i) = \begin{cases} \sqrt{(x_i(t) - x)^2 + (y_i(t) - y)^2}, & t - t_i \geq 0 \\ \infty, & \textit{otherwise}, \end{cases}$$

The anticipated position of pedestrian i at time t:
$x_i(t) = x_i(t_i) + (t - t_i)v_i^x(t_i), y_i(t) = y_i(t_i) + (t - t_i)v_i^y(t_i)$
The speed of pedestrian i at t_i in x and y directions: $v_i^x(t_i)$,$v_i^y(t_i)$

Mahalanobis distance (M-3DVoro)

$$d_M(p, p_i) = \sqrt{(p - p_i)^T M_i (p - p_i)}$$

M_i - symmetric, positive-definite matrix that defines how distances are measured from the perspective of pedestrian i

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

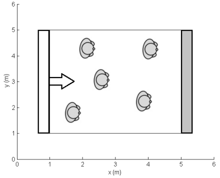
Performance of the approach

Synthetic data - unidirectional flow

NOMAD simulation tool [Campanella, 2010]

Scenario I: low congestion, homogenous population

Scenario II: high congestion, heterogeneous population



Indicators

Robustness w.r.t. the simulation noise

Robustness w.r.t. the sampling frequency

Characterization based on trajectories

Robustness with respect to the simulation noise

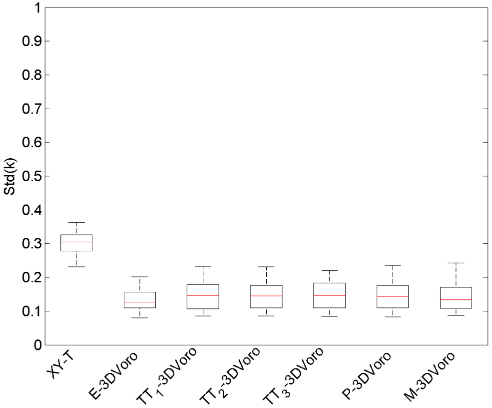
- 100 sets of pedestrian trajectories synthesized per scenario
- $\theta_r^M(p) = (k_r^M(p), v_r^M(p), q_r^M(p))$- a vector of indicators at point p obtained by applying the method M to the r^{th} set of trajectories
- The standard deviation of the indicators at p as

$$\sigma_R^M(p) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}(\theta_r^M(p) - \mu_R^M(p))^2}$$

$\mu_R^M(p) = \frac{1}{R}\sum_{r=1}^{R}\theta_r^M(p)$ and $R = 100$

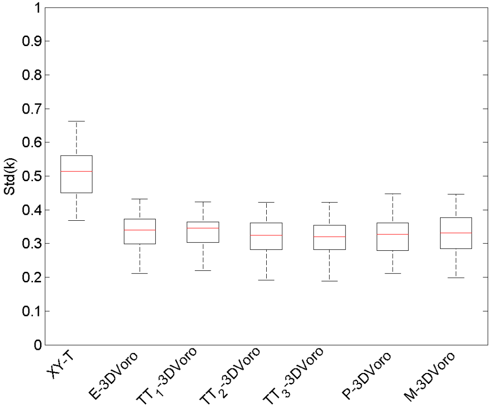
Robustness with respect to the simulation noise

Standard deviation (1000 points) - Scenario I



Robustness with respect to the simulation noise

Standard deviation (1000 points) - Scenario II



Characterization based on sampled data

Robustness with respect to the sampling frequency

- Ability of tolerating missing data

- Benchmark: indicators calculated on the true synthetic trajectories
- Sampled data: different sampling frequencies ($3s^{-1} - 0.5s^{-1}$)
- Indicators calculated via
  1. 3D Voro applied to the interpolated trajectories
  2. 3D Voro applied directly to the samples
- Comparison of the indicators to the corresponding benchmark values at 1000 randomly selected points

Robustness w.r.t the sampling frequency - Scenario I

High sampling frequency: $3s^{-1}$

Method	Mean		Mode		Median		90% quantile	
	IT	SoP	IT	SoP	IT	SoP	IT	SoP
XY-T	1.47E-02	/	1.25E-02	/	1.25E-02	/	6.25E-02	/
E-3DVoro	1.17E-02	/	0	/	4.48E-04	/	3.96E-02	/
TT_1-3DVoro	2.70E-03	6.70E-03	0	0	3.00E-04	2.30E-03	7.30E-03	1.02E-02
TT_2-3DVoro	5.80E-03	3.50E-02	0	2.80E-03	6.00E-04	2.08E-02	1.50E-02	6.69E-02
TT_3-3DVoro	5.40E-03	4.34E-02	0	8.00E-03	6.00E-04	2.83E-02	1.32E-02	9.22E-02
P-3DVoro	8.20E-03	5.36E-02	0	6.10E-03	2.40E-03	3.03E-02	1.30E-02	1.14E-01
M-3DVoro	4.50E-03	5.65E-02	0	6.80E-03	1.10E-03	4.55E-02	1.28E-02	1.04E-01

Low sampling frequency: $0.5s^{-1}$

Method	Mean		Mode		Median		90% quantile	
	IT	SoP	IT	SoP	IT	SoP	IT	SoP
XY-T	1.90E-01	/	1.00E-01	/	1.50E-01	/	3.38E-01	/
E-3DVoro	1.64E-01	/	1.12E-02	/	1.46E-01	/	3.02E-01	/
TT_1-3DVoro	2.54E-01	1.27E-01	1.35E-02	9.00E-03	1.16E-01	8.97E-02	3.41E-01	2.25E-01
TT_2-3DVoro	1.64E-01	1.22E-01	1.44E-02	1.06E-02	1.21E-01	7.30E-02	3.52E-01	2.33E-01
TT_3-3DVoro	1.89E-01	1.24E-01	1.84E-02	1.09E-02	1.24E-01	7.88E-02	3.40E-01	2.31E-01
P-3DVoro	3.19E-01	1.21E-01	3.26E-02	6.20E-03	1.43E-01	7.43E-02	3.36E-01	2.10E-01
M-3DVoro	1.97E-01	1.24E-01	3.48E-02	9.90E-03	1.41E-01	7.72E-02	3.21E-01	2.31E-01

Robustness w.r.t the sampling frequency - Scenario II

High sampling frequency: $3s^{-1}$

Method	Mean		Mode		Median		90% quantile	
	IT	SoP	IT	SoP	IT	SoP	IT	SoP
XY-T	2.05E-02	/	0	/	1.25E-02	/	5.00E-02	/
E-3DVoro	1.43E-02	/	0	/	2.67E-02	/	2.64E-02	/
TT_1-3DVoro	8.00E-03	4.55E-02	0	0	8.00E-04	1.75E-02	2.36E-02	8.52E-02
TT_2-3DVoro	1.49E-02	1.07E-01	0	0	3.20E-03	5.72E-02	3.33E-02	2.21E-01
TT_3-3DVoro	1.24E-02	1.60E-01	0	0	3.50E-03	9.62E-02	2.98E-02	3.41E-01
P-3DVoro	2.10E-02	1.66E-01	0	0	4.20E-03	1.16E-01	5.27E-02	3.64E-01
M-3DVoro	1.31E-02	2.40E-01	0	0	2.50E-03	1.75E-01	2.91E-02	5.58E-01

Low sampling frequency: $0.5s^{-1}$

Method	Mean		Mode		Median		90% quantile	
	IT	SoP	IT	SoP	IT	SoP	IT	SoP
XY-T	5.29E-01	/	1.63E-01	/	4.75E-01	/	1.01E00	/
E-3DVoro	4.02E-01	/	0	/	2.49E-01	/	1.03E+00	/
TT_1-3DVoro	4.06E-01	2.90E-01	3.10E-01	2.48E-02	2.64E-01	1.65E-01	9.21E-01	7.12E-01
TT_2-3DVoro	3.92E-01	4.58E-01	2.85E-01	2.34E-01	2.48E-01	2.34E-01	9.30E-01	1.11E+00
TT_2-3DVoro	4.41E-01	5.07E-01	2.89E-01	5.89E-02	2.37E-01	3.06E-01	9.81E-01	1.17E+00
P-3DVoro	4.31E-01	3.71E-01	1.40E-03	0	2.58E-01	1.80E-01	9.43E-01	7.29E-01
M-3DVoro	4.34E-01	5.01E-01	3.16E-01	1.36E-01	2.75E-01	3.52E-01	9.96E-01	9.80E-01

Outline

1. Introduction
2. Related research
3. Methodology
   - Discretization framework
   - Definitions of the indicators
   - Spatio-temporal distances
4. Empirical analysis
5. Conclusion and future work

Conclusion

- A novel approach to pedestrian traffic characterization: data-driven discretization via 3D Voronoi diagrams
- Discretization based on trajectories available either in the form of an analytical description or as a finite collection of points
- The exact characterization of the Voronoi diagrams can be adapted to specific situation
- Superior to existing methods w.r.t. robustness to the simulation noise
- Robustness to the sampling frequency
  - Higher sampling frequency: 3DVoro based on interpolated trajectories shows better results (Time-Transform 3D Voronoi)
  - Lower sampling frequency: 3DVoro based on sample of points exhibit better performance (anticipating distances)

Future work

- Analysis of the performance for other behavioral situations (bi-directional and multi-directional scenarios)
- The effectiveness of the approach using real data (railway station, Lausanne)
- Characterization in the presence of obstacles
- Weighted assignment rules to account for the anisotropy of pedestrian movements

Thank you

hEART 2016 - 5^{th} Symposium of the European Association for Research in Transportation, Delft University of Technology:

Data-driven characterization of pedestrian traffic

Marija Nikolić, Michel Bierlaire

Help by S. S. Azadeh and F.Hänseler is appreciated.

- marija.nikolic@epfl.ch